Wavelets and WMAP non-Gaussianity

Article (Published Version)

The Astrophysical Journal, 613:51–60, 2004 September 20

WAVELETS AND *WILKINSON MICROWAVE ANISOTROPY PROBE* NON-GAUSSIANITY

Pia Mukherjee and Yun Wang
Department of Physics and Astronomy, University of Oklahoma, 440 West Brooks Street, Norman, OK 73019;
pia@nhn.ou.edu, wang@nhn.ou.edu
Received 2004 February 25; accepted 2004 June 2

ABSTRACT

We study the statistical properties of the first-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*) data on different scales using the spherical Mexican hat wavelet transform. Consistent with the 2004 results of Vielva et al., we find a deviation from Gaussianity in the form of kurtosis of wavelet coefficients on 3°–4° scales in the southern Galactic hemisphere. This paper extends the work of Vielva et al. as follows: We find that the non-Gaussian signal shows up more strongly in the form of a larger than expected number of cold pixels and also in the form of scale-scale correlations among wavelet coefficients. We establish the robustness of the non-Gaussian signal under more wide-ranging assumptions regarding the Galactic mask applied to the data and the noise statistics. This signal is unlikely to be due to the usual quadratic term parameterized by the nonlinearity parameter $f_{\rm NL}$. We use the skewness of the spherical Mexican hat wavelet coefficients to constrain $f_{\rm NL}$ with the first-year *WMAP* data. Our results constrain $f_{\rm NL}$ to be 50 ± 80 at 68% confidence and less than 280 at 99% confidence.

Subject headings: cosmic microwave background — methods: statistical

1. INTRODUCTION

The current cosmological model assumes Gaussian initial conditions, created by inflation. This assumption regarding the nature of primordial density perturbations can be verified by studying the distribution of temperature fluctuations in the cosmic microwave background (CMB). While the simplest inflationary models predict Gaussian primordial perturbations, there are other models of inflation, such as those involving multiple scalar fields, features in the inflation potential, or phase transitions, that could give rise to detectable non-Gaussianity. Hence, studies of Gaussianity help distinguish between different early universe scenarios. Gaussianity is also a key underlying assumption of CMB data analysis wherein the angular power spectrum fully specifies its statistical properties and must be tested. Non-Gaussianity can also be associated with secondary anisotropies in the CMB or with foreground contamination and systematic effects.

Prior to the release of *Wilkinson Microwave Anisotropy Probe* (*WMAP*) data there was no clear evidence of cosmological non-Gaussianity. Since the release of the first-year *WMAP* data, a number of tests of non-Gaussianity have been performed, with somewhat differing results. Each statistic is sensitive to a different kind of non-Gaussianity; hence, there is a need for a wide variety of tests. Komatsu et al. (2003b) use an optimized test based on the bispectrum, as well as Minkowski functionals, while Colley & Gott (2003) study the genus, and both groups report consistency with Gaussianity. Gaztanaga & Wagg (2003) do a three-point angular correlation function analysis and find consistency with Gaussianity as well. Chiang et al. (2003) perform a study of the phases of spherical harmonics and find some evidence for non-Gaussianity at high multipoles. Copi et al. (2004) find some evidence for low-l correlations and deviation from isotropy. Park (2004) finds a large difference between the genus amplitudes of the northern and southern hemispheres and a positive genus asymmetry in the southern hemisphere. Eriksen et al. (2004b) compute the two- and three-point correlations and report a significant north-south asymmetry; Eriksen et al. (2004a) use Minkowski functionals and find a significant genus in the northern hemisphere and, again, indications of north-south asymmetry. Hansen et al. (2004) use local curvature and find non-Gaussianity/asymmetry in the data on scales of a few degrees. Gurzadyan et al. (2004) find ellipticity in the temperature anisotropy features in the data, consistent with what was found previously in BOOMERANG data. Vielva et al. (2004, hereafter V04) report a non-Gaussian signal in the southern hemisphere at high significance in the form of kurtosis on $\sim 4^\circ$ scales using the spherical Mexican hat wavelet (SMHW) transform on *WMAP* data. Some of the detections of non-Gaussianity and/or asymmetry thus far reported in the *WMAP* data are at the level of 99% or greater.

Wavelet transforms are useful tools in non-Gaussianity studies because they enable the signal on the sky to be studied on different scales, with simultaneous position localization, so that the obscuring effects of the central-limit theorem, which can exist in both real and Fourier spaces, are reduced. With wavelets, any non-Gaussian detection can be localized on the sky in scale and position, so that its nature and source can be better determined. Planar wavelets have been used in Gaussianity studies of the CMB by Pando et al. (1999), Hobson et al. (1999), and Mukherjee et al. (2000), while Barreiro et al. (2000) use the spherical Haar wavelet, and Cayón et al. (2001, 2003), Martínez-González et al. (2002), and V04 use the SMHW. Wavelet methods have been compared with other pixel or Fourier-based methods in Hobson et al. (1999), Aghanim et al. (2003), and Cabella et al. (2004), and the performance of isotropic, as well as highly anisotropic, multiscale bases in distinguishing between different sources of non-Gaussianity in the CMB has been studied in Starck et al. (2004).

In this paper we use the SMHW transform to probe non-Gaussianity in the *WMAP* data. We extend the work of V04 by performing new multiple tests of the robustness of the non-Gaussian signal. We also look at the non-Gaussianity in terms of an excess in the number of cold pixels and in terms of scale-scale correlations among wavelet coefficients. Furthermore, we place constraints on a popular form of non-Gaussianity (a quadratic term in the curvature perturbations parameterized by

TABLE 1
Scales Used in the Spherical Mexican Hat Wavelet Transform

Scale	Size (arcmin)
R_1	14
R_2	25
R_3	50
R_4	75
R_5	100
R_6	150
R_7	200
R_8	250
R_9	300
R_{10}	400
R_{11}	500
R_{12}	600
R_{13}	750
R_{14}	900
R_{15}	1050

the nonlinearity parameter $f_{\rm NL}$). This paper is organized as follows: In § 2 we present results from using the SMHW transform on *WMAP* data. While confirming the results of V04, we perform new multiple tests of the robustness of the non-Gaussianity signal in the kurtosis spectrum (1) through the use of different (extended) masks and (2) by relaxing the assumption of a simplified noise model. We find that the signal shows up even more significantly in the form of the number of cold pixels (or coefficients). In § 3 we examine scale-scale correlations among the wavelet coefficients. We find significant deviations from Gaussianity, a corroboration of the signal detected and described in § 2. In § 4 we obtain constraints on the nonlinearity parameter $f_{\rm NL}$. We conclude in § 5.

2. SKEWNESS AND KURTOSIS OF WAVELET COEFFICIENTS

A non-Gaussianity detection in the first-year *WMAP* data was reported by V04. Applying the SMHW transform to the Q-V-W co-added data and computing the skewness and kurtosis of the wavelet coefficients over scales ranging from about $10'$ to 10°, they found that the kurtosis of wavelet coefficients on scales of $\sim 4^\circ$ was too high at the 99.9% confidence level. It was found that the excess kurtosis was in the southern hemisphere, while the kurtosis signal in the northern hemisphere was consistent with Gaussianity. The signal was shown to be independent of frequency.

It is important to determine whether the significance of this non-Gaussianity detection is influenced by systematic effects, such as the choice of mask, or simplified assumptions about noise. Since the SMHW transform is a sensitive probe of non-Gaussianity, we use it to perform an independent analysis of the first-year *WMAP* data. The basic steps followed in the analysis are as follows: starting with the foreground cleaned Q-V-W co-added data, we bring the map down in resolution to HEALPix $n_{\rm side} = 256$, apply the Kp0[1] mask, perform SMHW transforms (see the Appendix) to obtain wavelet coefficients corresponding to the different scales R (also setting the monopole and dipole of the map to zero here), apply appropriately extended versions of the mask to the wavelet coefficients of each scale to exclude coefficients contaminated by the mask and known point sources, and compute the skewness and kurtosis of the remaining unmasked coefficients.

For our results to be directly comparable we perform the SMHW analysis for the same scales used by V04. For convenience and clarity, the scales R_i ($i = 1, 2, \ldots, 15$) plotted in the subsequent figures are listed in Table 1.

Figure 1 shows our results, with the extended mask at each scale obtained by extending the Kp 0+sources mask such that all pixels closer than $2.5R$ to any of the pixels in the Kp0+sources mask within $|b| < 25^\circ$ are excluded from the analysis, attempting to follow closely the procedure of V04. The mask was not extended around point sources outside of this region, as the Kp0+sources mask around point sources seems to be in general extended enough to not cause contamination in wavelet coefficients on small scales, and on larger scales the effect gets averaged out.[2] Conservatively extending the region around point sources would also leave too few pixels on scales of interest here. (It is unlikely that the non-Gaussianity signal found below, on $\sim 4^\circ$ scales, is coming from point sources; see also Fig. 4*b* and related text in § 2.1).

The mean and 1, 2, and 3 σ confidence contours obtained from 1000 Gaussian simulations processed in the same way as the data are also shown in Figure 1. The Gaussian simulations were created in the following way: CMB realizations with the same C_l spectrum as the flat-Λ cold dark matter cosmology with a power-law primordial power spectrum that best fits the first-year *WMAP* data (Spergel et al. 2003; Hinshaw et al. 2003) were created at HEALPix resolution $n_{\rm side} = 512$. Each realization was copied and smoothed with the *WMAP* beam window functions for each of the Q, V, and W radiometer channels. Independent noise realizations of rms $\sigma_0/N_{\rm obs}^{1/2}$ were then added to the maps, where the effective number of observations $N_{\rm obs}$ varies across the sky and σ_0 is different for each radiometer channel. The eight maps thus produced were co-added, weighted by $N_{\rm obs}/\sigma_0^2$. Thereafter, the same analysis procedure that was applied to the data map was applied to each of the Gaussian simulations.

If we estimate the significance of the signal using the generic χ^2 test, which includes information on all scales,

$$\chi^2 = \sum_{R_i, R_j} [S(R_i) - \bar{S}(R_i)] \Sigma^{-1}_{R_i, R_j} [S(R_j) - \bar{S}(R_j)], \qquad (1)$$

where $S(R_i)$ is the skewness or kurtosis signal on scale R_i, $\bar{S}(R_i)$ is the mean obtained from Gaussian simulations, and Σ_{R_i, R_j} is the covariance matrix obtained from simulations, we find that by comparing the χ^2 of the signal in the data with the distribution of χ^2 values obtained from Gaussian realizations, the kurtosis signals in nine of 1000 realizations have larger χ^2 values than the data. Hence, we arrive at a significance of 99% for this signal. The skewness signal is consistent with Gaussianity. Significances obtained using the χ^2 test are tabulated in Table 2, under mask 1.

[1] By Kp0 we mean the Kp0 mask without sources.

[2] There are, however, three bright semipoint sources in the northern hemisphere that are seen to visibly contaminate wavelet coefficients on scales less than R_7. These are taken care of by extending the mask out to $2.5R$ around these sources for scales less than R_7. There is also a more diffuse spot in the Kp0 mask outside the $|b| < 25^\circ$ region that causes visible contamination, even on larger scales. The mask around this region is extended as well. The extended masks around these four regions are always retained, even when the remaining details of the mask are varied. For reference, these regions are shown mapped in Fig. 2 of Eriksen et al. (2004b). The actual number of pixels in the co-added map that contain emission from these four sources together is quite small (<40, for $n_{\rm side} = 256$).

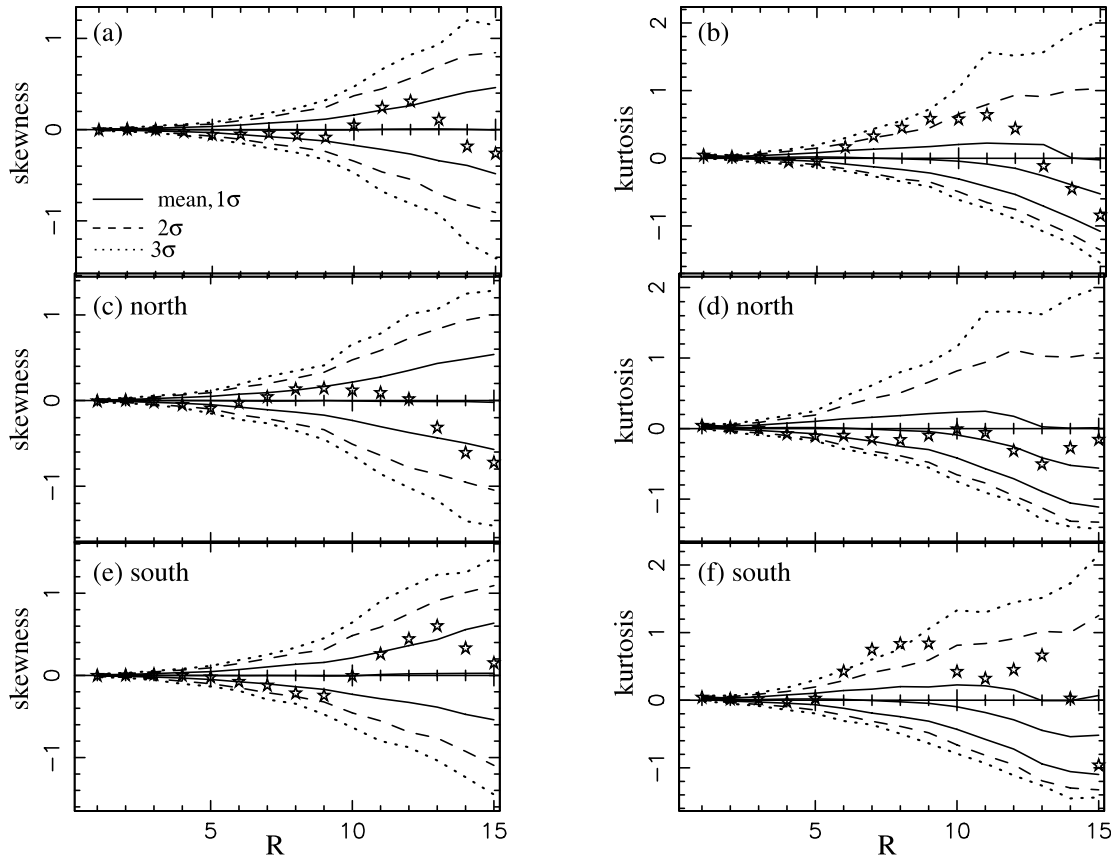

FIG. 1.—Skewness and kurtosis spectra obtained using a SMHW analysis of the *WMAP* co-added data (*stars*), together with the mean values, and 1 (*solid lines*), 2 (*dashed lines*), and 3 σ (*dotted lines*) confidence contours obtained from 1000 Gaussian simulations. The top panels show these for all sky, the middle for the northern Galactic hemisphere, and the bottom for the southern Galactic hemisphere. Similar results were presented in V04.

We find good consistency with the signal reported in V04. The kurtosis signal on 3°–4° scales in the southern Galactic hemisphere is outside the 3 σ confidence contour. Only two and three simulations out of 1000 lead to a stronger kurtosis signal on scales R_7 and R_8, respectively, and only one of the Gaussian realizations has a larger kurtosis than the data on both the R_7 and R_8 scales in the southern hemisphere. Hence, the signal on these particular scales appears significant at the 99.9% confidence level.

Even if the data were Gaussian, there would be a certain probability of obtaining outlier signals in its kurtosis spectrum on at least two of the 15 scales considered. We take a closer look at this probability to better understand the significance of the detection. The number of Gaussian realizations that have kurtosis values that fall outside the 99% confidence region in the southern hemisphere in *any* two of the 15 scales is 28 (this number is 17 for a positive kurtosis), indicating that the signal found above on scales R_7 and R_8 in the data is significant at *at least* the 97% confidence level. (The two scales were always consecutive in the simulations but well spread out among all the scales.) Furthermore, these numbers are 22 and 17 for the northern hemisphere, and in none of these cases did the signals on both hemispheres lie outside their 99% limits, implying that a significant north-south asymmetry in the kurtosis signal in *any* two scales was seen in 50 (37 for positive kurtosis) of 1000 Gaussian realizations. Thus, the north-south asymmetry itself appears significant at *at least* the 95% level.

TABLE 2

SIGNIFICANCE OF DEVIATION OF THE SKEWNESS AND KURTOSIS SIGNALS FROM GAUSSIANITY USING THE χ^2 TEST

Mask	Skewness (%)	Kurtosis (%)
Mask 1	44	99
Mask 2	13	99
Mask 3	57	98
Mask 4	79	99
ILC	24	95

Histogram plots of the wavelet coefficients on scale R_7 are shown in Figure 2, for the all-sky, northern, and southern Galactic hemispheres. A longish tail toward negative values is seen in the southern hemisphere. Figure 3 (*top panels*) shows statistics relating to the minima, maxima (both in units of σ on each scale), and σ of the wavelet coefficients in the southern hemisphere, plotted against scale. We use σ to denote the rms dispersion of the wavelet coefficients on each scale, noting that 1 σ, 2 σ, etc., may not correspond to the same confidence levels as for a Gaussian distribution. The middle and bottom panels show the statistics relating to the number of wavelet coefficients that were larger than mean $+1$ σ, mean $+2$ σ, and mean $+3$ σ and smaller than mean -1 σ, mean -2 σ, and mean -3 σ, respectively, again in the southern hemisphere. While the rest are seen to be quite consistent with limits obtained from Gaussian simulations, we see that the minima on $\sim 4^\circ$ scales is significant (with only $\sim$1% of the simulations

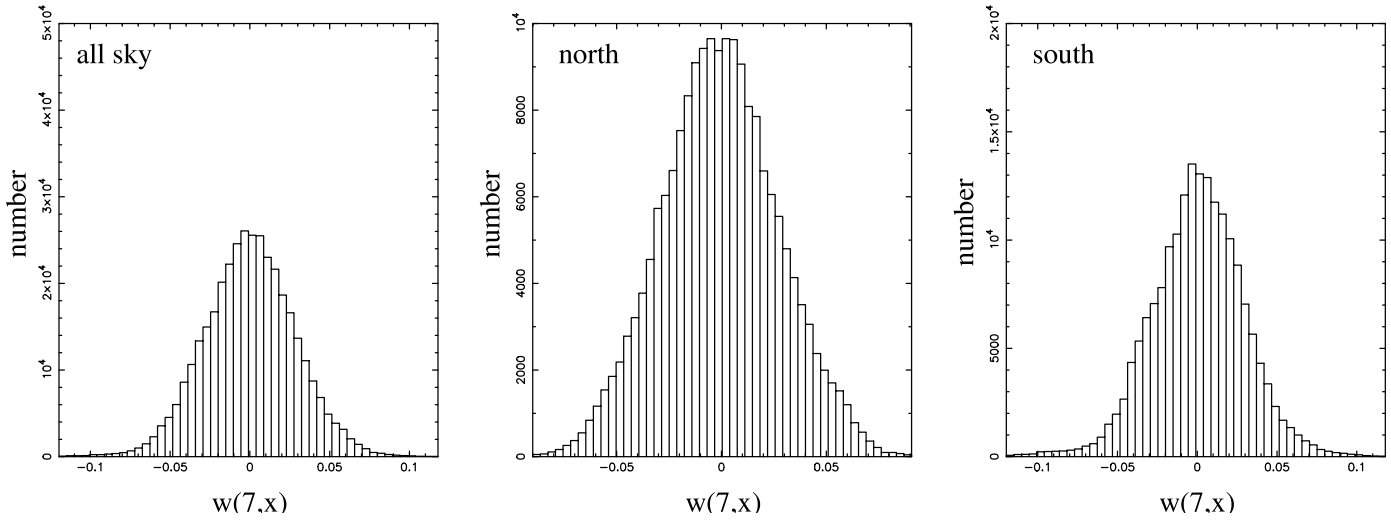

FIG. 2.—Histograms of wavelet coefficients on scale R_7, for the all-sky, northern, and southern Galactic hemispheres.

showing a stronger minima on each of scales R_8 and R_9), and the number of wavelet coefficients that are smaller than mean $-3\ \sigma$ is very significant, with none of the simulations showing a stronger deviation on scales R_6 and R_7. This last estimator clearly gives a very strong signal. Only three of 1000 Gaussian realizations give a value for this estimator that is larger than the 99% confidence contour in *any* four scales. The signal we have here thus appears more significant than 99.7%, as in the data the value of this estimator on four scales is well out of the 99% confidence region. The χ^2 test described earlier also gives a similar significance for this signal. The number of cold pixels in the southern hemisphere is too large, on scales of 3°–4°. On scales R_6 and R_7 there is more than one cold spot contributing to this number, while on larger scales it is mainly the one cold spot located near ($b = -57^\circ$, $l = 209^\circ$) pointed out in V04. This spot is present on scales R_6 and R_7

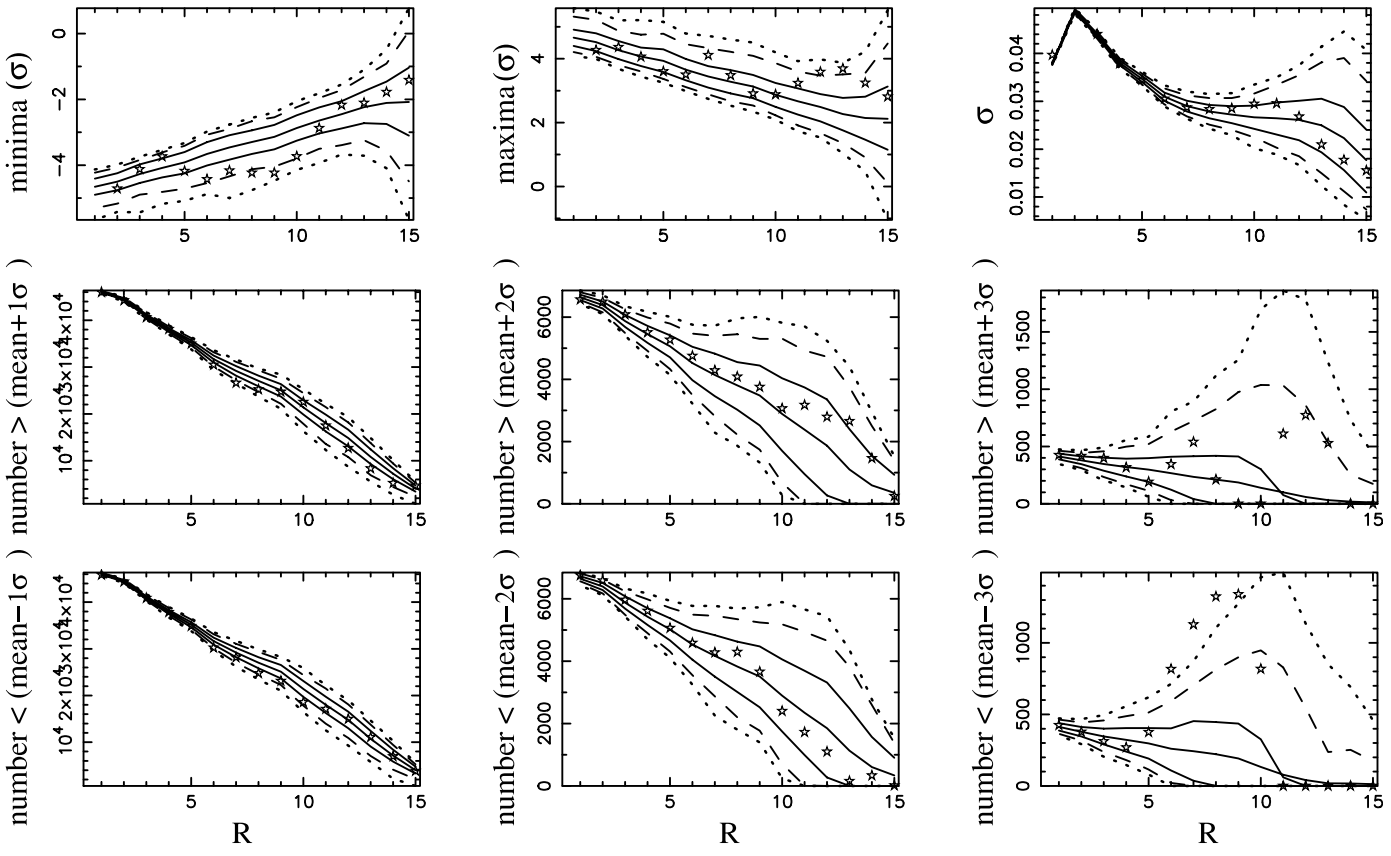

FIG. 3.—Statistics relating to the minima and maxima, in units of σ, and σ itself (*top panels*) and the number of wavelet coefficients that are larger than mean $+1\ \sigma$, mean $+2\ \sigma$, and mean $+3\ \sigma$ (*middle panels*) and smaller than mean $-1\ \sigma$, mean $-2\ \sigma$, and mean $-3\ \sigma$ (*bottom panels*), all for the southern Galactic hemisphere. When comparing the data (co-added *WMAP*; *stars*) with the confidence limits obtained from Gaussian simulations (mean and 1 [*solid lines*], 2 [*dashed lines*], and 3 σ [*dotted lines*]), the data show an excess of cold coefficients.

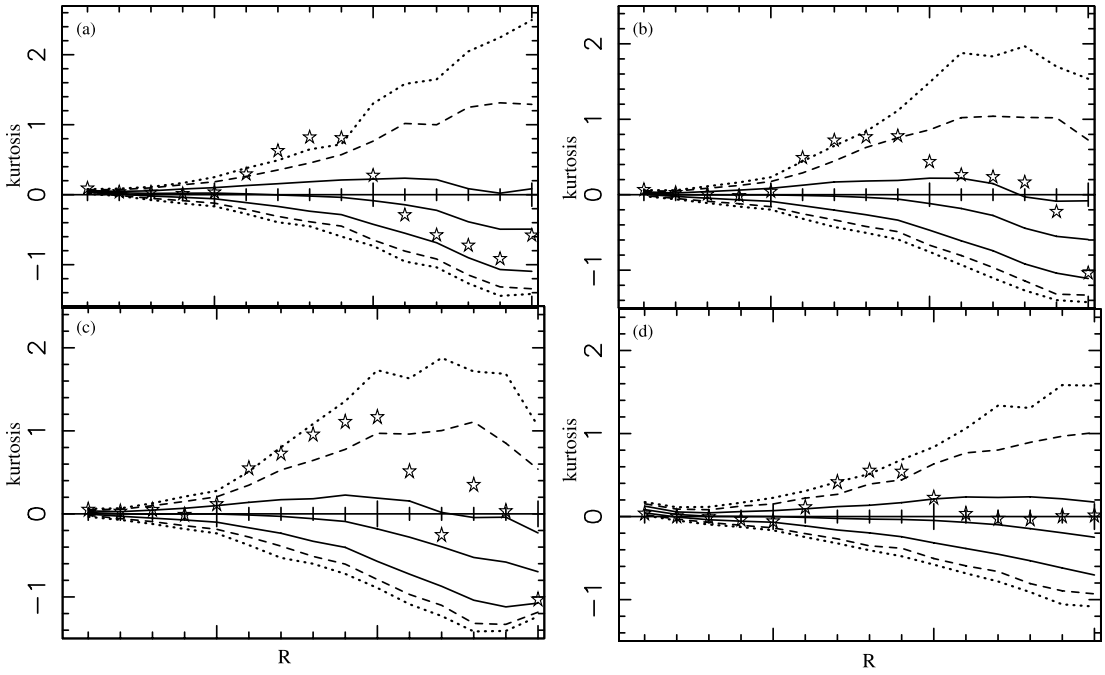

FIG. 4.—Kurtosis spectra of wavelet coefficients in the southern Galactic hemisphere for different masks, as described in the text.

as well. The corresponding statistics for the northern hemisphere are very consistent with Gaussianity. Hence, the non-Gaussianity shows up in the southern hemisphere in the form of a kurtosis signal and a larger than expected number of cold pixels.

2.1. *Other Masks*

We have checked that the above results are unaltered on using a variety of different masks. We now show the kurtosis signal in the southern hemisphere for a few different masks. The mean, 1, 2, and 3 σ confidence contours are obtained from Gaussian simulations processed each time in the same way as the data.

Figure 4*a* shows the result of using an extended mask that is made in the same way as for Figure 1 above but by extending the boundaries of the whole Kp 0 (without sources) mask by $2.5R$ and then adding the mask around point sources back in on scales smaller than R_7. The shape of the mask is then retained on all scales.[3]

In Figure 4*b* we show the result of using extended masks that on each scale apply a straight $|b| < (25 + 2.5R)^\circ$ Galactic cut, as well as a mask around point sources for scales smaller than R_7. The signal is thus unaffected by the shape of the mask. Figure 4*c* shows the result of using a straight $|b| < (35 + 2.5R)^\circ$ Galactic cut. In going from Figure 4*a* to 4*c* more of the sky is being excluded by the extended mask, and it is seen in the form of increased variance (this effect being larger for larger scales). But while the significance of the kurtosis signal seems to go down in this way, Figure 5 shows that the number of cold pixels are in fact equally or more significant for this last mask.

Figure 4*d* shows the results from the internal linear combination (ILC) map. In this case, since there is little contamination from the Galactic plane left in the map, we can use just the Kp0+sources mask without any extensions and apply it after the SMHW transform. The Gaussian simulations here were obtained by simulating the signal in each of the 10 radiometer channels, smoothing to 1° resolution, obtaining the noise-weighted averaged signal for the five frequency channels, and then taking a linear combination of these with the weights given in § 4 of Bennett et al. (2003). The signal is found in the ILC map as well.

From the above analysis it is found that the kurtosis signal is indeed independent of the properties of the mask. The significances derived using the χ^2 test for above masks are in Table 2. The masks corresponding to Figures 4*a*, 4*b*, and 4*c* are labeled "mask 2," "mask 3," and "mask 4," respectively, in the table; the ILC case with the unextended Kp0+sources mask is labeled "ILC."

2.2. *Noise Simulations*

Finally, Figure 6 shows the kurtosis signal for the all-sky case for the same mask as used in Figure 1, but this time with confidence contours obtained from 110 Gaussian simulations that make use of the 110 full-noise simulations provided by the *WMAP* team for each radiometer channel. The full-noise simulations were made by generating 1 year of simulated time-series data including white noise, $1/f$ noise, and all interchannel correlations that are known to exist in the radiometers, and then taking this data through all the steps of processing, such as flight calibration, map making, and pipeline filtering. We found that when compared with the case of 110 realizations of simple white noise for the same underlying sky simulations, the two cases give identical results, so that when plotted simultaneously they are indistinguishable. This indicates that our simulations are reliable and that the simple

[3] We have also checked that applying the Kp0 mask before or after performing the SMHW transform (but before applying the extended masks) does not affect the signal. In addition, if we apply the SMHW transform to the Kp0 mask, identify the wavelet coefficients that are, for instance, more than 1% contaminated (by the edge of the mask), and include these pixels to make the extended masks, we check that the signal remains unaffected.

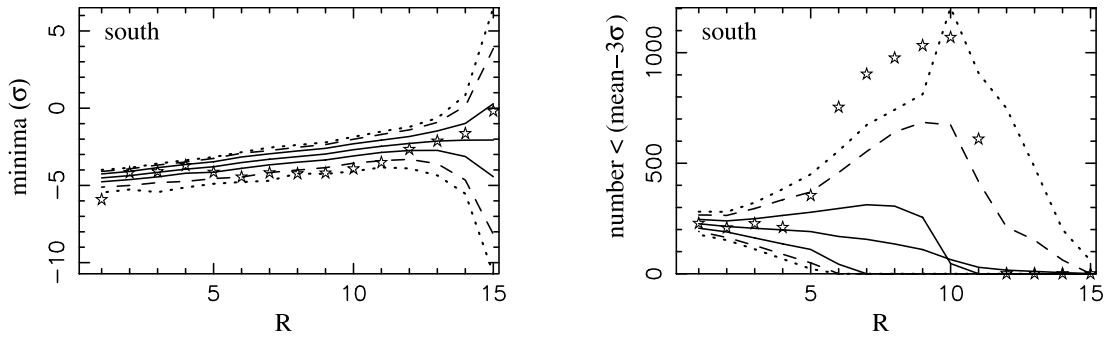

FIG. 5.—Statistics relating to the minima, shown in units of σ, and to the number of wavelet coefficients that are smaller than mean -3 σ in the *WMAP* co-added data (*stars*) in the southern Galactic hemisphere for a mask that is a straight $|b| < (35 + 2.5R)^{\circ}$ Galactic cut (see text). Also shown are the mean and 1 (*solid lines*), 2, (*dashed lines*), and 3 σ (*dotted lines*) confidence contours obtained from Gaussian realizations.

white noise model is completely satisfactory according to this statistic.

3. SCALE-SCALE CORRELATIONS

Having obtained the wavelet coefficients of the data on different scales, we compute the scale-scale correlations between corresponding coefficients,

$$C_{R_i,R_j} = \frac{N\sum_x w(R_i,x)^2 w\left(R_j,x\right)^2}{\sum_x w(R_i,x)^2 \sum_x w\left(R_j,x\right)^2}, \qquad (2)$$

where $w(R_i, x)$ are wavelet coefficients on scale R_i and at position or pixel x in the sky. The coefficients that contribute to the sums are the N unmasked coefficients on the larger scale.

Figure 7 (*top*) shows scale-scale correlations between wavelet coefficients. The exact order of plotting is given in Table 3. Coefficients over the whole sky are used here. The mean and 1, 2, and 3 σ confidence contours obtained for the scale-scale correlations from Gaussian simulations are also shown. Consistency with Gaussianity is indicated among these well-separated scales.

Zooming into the scales that indicated non-Gaussianity in § 2, Figure 7 (*bottom*) shows the scale-scale correlations between coefficients of scales R_6–R_{11} in the southern Galactic hemisphere. The exact order of plotting is given in Table 3. Compared with Gaussian simulations, we now find significant scale-scale correlations between the wavelet coefficients, presumably due to the cold spot that is common to all of these scales. A similar result is found for different masks.

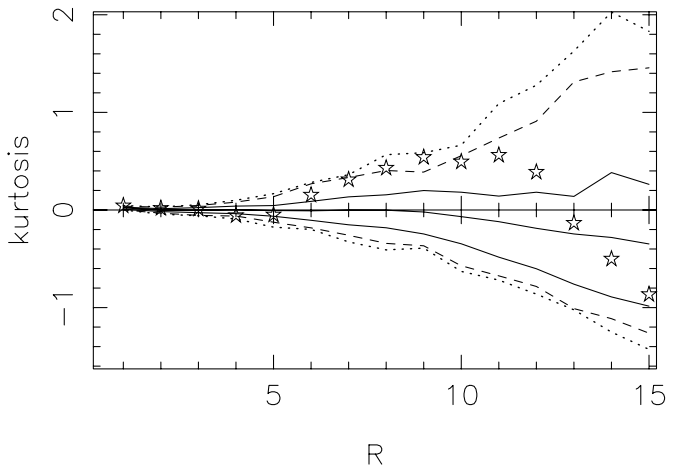

FIG. 6.—Plot to be compared with Fig. 1*b*. The confidence contours here have been obtained from 110 simulations using the 110 full-noise simulations for each radiometer channel made available by the *WMAP* team.

4. $f_{\rm NL}$ CONSTRAINTS

We showed in § 2 that the skewness signal obtained using a range of masks was consistent with Gaussianity. In this section we use the skewness signal on the 15 scales to compute the limits that can be placed on the amplitude of primordial non-Gaussianity, as parameterized by the nonlinear coupling parameter $f_{\rm NL}$. Since the nonlinear term (defined below) is not dominant and is quadratic, it is the skewness signal in the data, rather than any higher order cumulants, that is expected to be most sensitive to it. The kurtosis signal is not sensitive to this kind of non-Gaussianity (Cayón et al. 2003).

The quantity $f_{\rm NL}$ characterizes the amplitude of a quadratic term added to the curvature perturbations,

$$\Phi(x) = \Phi_L(x) + f_{\rm NL}\left[\Phi_L^2(x) - \langle\Phi_L^2(x)\rangle\right], \qquad (3)$$

where Φ_L are Gaussian linear perturbations with zero mean. Thus, $f_{\rm NL}$ parameterizes the leading order nonlinear corrections to the primordial (curvature) perturbations. The motivation to use data to place constraints on $f_{\rm NL}$ is to address how Gaussian the current CMB data are or how much primordial non-Gaussianity of this particular form is allowed by the data. Such analyses also help compare the sensitivity of different data sets and of different estimators of non-Gaussianity to this particular form of non-Gaussianity.

Using an optimal estimator of non-Gaussianity based on the bispectrum, namely, the cubic statistic, Komatsu et al. (2003b) place limits of $-58 < f_{\rm NL} < 134$ at 95% confidence using the first-year *WMAP* data. They derive 95% confidence limits of $f_{\rm NL} < 139$ based on Minkowski functionals. Smith et al. (2004), using Very Small Array (VSA) data, obtain an upper limit of 3100 at 95% confidence (their limit being large because their data are sensitive to small scales). Previously, using *COBE* DMR data, the bispectrum analysis of Komatsu et al. (2002) placed a limit of $|f_{\rm NL}| < 1500$, and using the skewness of SMHW coefficients, Cayón et al. (2003) placed a limit of $|f_{\rm NL}| < 1100$, both at 68% confidence. Using MAXIMA data, Santos et al. (2003) placed a 1 σ limit of $|f_{\rm NL}| < 950$.

We use the non-Gaussian simulations of Komatsu et al. (2003b). How these were produced is described in detail in the Appendix of their paper. Since producing these maps is a computationally intensive process, we use the 300 available realizations of non-Gaussian sky maps at HEALPix resolution

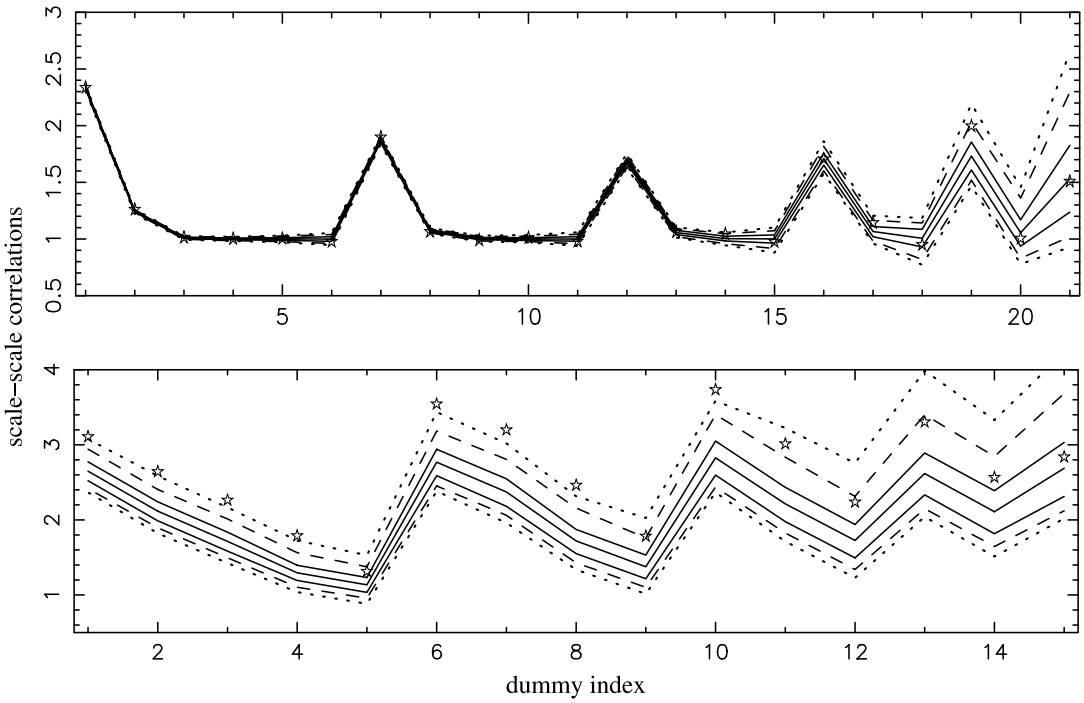

FIG. 7.—*Top*: Scale-scale correlations among wavelet coefficients between scales that are well separated and span the whole range studied here. *Bottom*: Scale-scale correlations among wavelet coefficients in the southern Galactic hemisphere between scales R_6 and R_{11}, the scales that show excess kurtosis and an excess of cold coefficients. See Table 3 for the exact order of plotting. The correlations obtained from the *WMAP* co-added data are shown as stars. Also shown are the mean and 1 (*solid lines*), 2 (*dashed lines*), and 3 σ (*dotted lines*) confidence contours obtained from Gaussian realizations.

$n_{\rm side} = 256$ to obtain the mean values of skewness at each of the different scales, for different values of $f_{\rm NL}$. We use Gaussian simulations to estimate the covariance matrix of the skewness values for the different scales and thus in turn to estimate the uncertainty in the measured $f_{\rm NL}$. Since Gaussian simulations can be computed several orders of magnitude in time faster, we can estimate the covariance matrix accurately using a larger number of Gaussian simulations, and the uncertainty estimated from Gaussian simulations is a good approximation to that estimated from non-Gaussian simulations for $|f_{\rm NL}| < 500$.[4]

[4] One way to see this is that the 1 σ uncertainties indicated by the different curves in Fig. 8 below are roughly the same, indicating that at the current sensitivity of the data, the uncertainties do not depend on the precise best-fit value of $f_{\rm NL}$.

In order to estimate the maximum likelihood $f_{\rm NL}$, we compare the skewness values of the data at the 15 scales with simulations and use the goodness-of-fit statistic

$$\chi^2 = \sum_{R_i, R_j} \left[S(R_i) - \bar{S}_{\rm sim}(R_i)\right] \Sigma^{-1}_{R_i, R_j} \left[S(R_j) - \bar{S}_{\rm sim}(R_j)\right], \quad (4)$$

where $S(R_i)$ is the skewness of the *WMAP* data on the ith scale R_i, $\bar{S}_{\rm sim}(R_i)$ is the mean value from Monte Carlo simulations, computed for different values of $f_{\rm NL}$, and Σ_{R_i, R_j} is the scale-scale skewness covariance matrix from simulations.

We have tested that the χ^2 statistic accurately recovers $f_{\rm NL}$ by using it on simulated maps. This is illustrated in Figure 8. The figure shows the mean χ^2 distributions obtained from 300 simulated realizations of non-Gaussian maps with $f_{\rm NL}$ values

TABLE 3
SCALES ASSOCIATED WITH THE DUMMY INDICES PLOTTED IN FIG. 7

TOP PANEL				BOTTOM PANEL			
Index	Scales	Index	Scales	Index	Scales	Index	Scales
1	R_1, R_2	12	R_3, R_5	1	R_6, R_7	12	R_8, R_{11}
2	R_1, R_3	13	R_3, R_7	2	R_6, R_8	13	R_9, R_{10}
3	R_1, R_5	14	R_3, R_{10}	3	R_6, R_9	14	R_9, R_{11}
4	R_1, R_7	15	R_3, R_{14}	4	R_6, R_{10}	15	R_{10}, R_{11}
5	R_1, R_{10}	16	R_5, R_7	5	R_6, R_{11}		
6	R_1, R_{14}	17	R_5, R_{10}	6	R_7, R_8		
7	R_2, R_3	18	R_5, R_{14}	7	R_7, R_9		
8	R_2, R_5	19	R_7, R_{10}	8	R_7, R_{10}		
9	R_2, R_7	20	R_7, R_{14}	9	R_7, R_{11}		
10	R_2, R_{10}	21	R_{10}, R_{14}	10	R_8, R_9		
11	R_2, R_{14}			11	R_8, R_{10}		

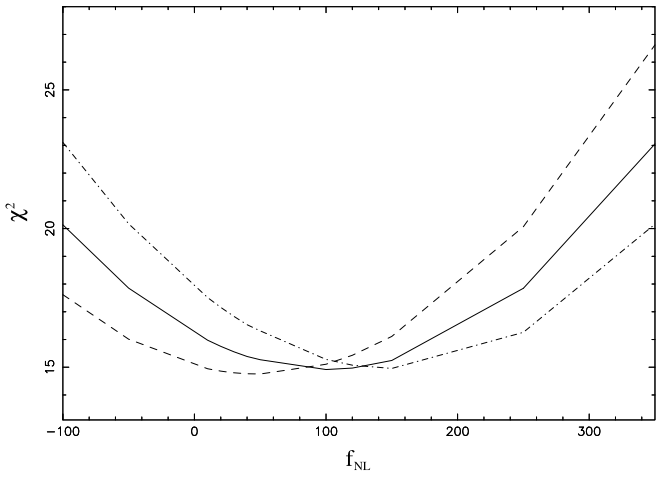

FIG. 8.—Mean χ^2 distributions obtained from 300 simulated realizations of non-Gaussian maps with $f_{\rm NL}$ values of 50 (*dashed lines*), 100 (*solid lines*), and 150 (*dot-dashed lines*), illustrating that the statistic recovers the inputted value of $f_{\rm NL}$. The simulations include, as usual, the noise properties and window functions of the first-year *WMAP* data.

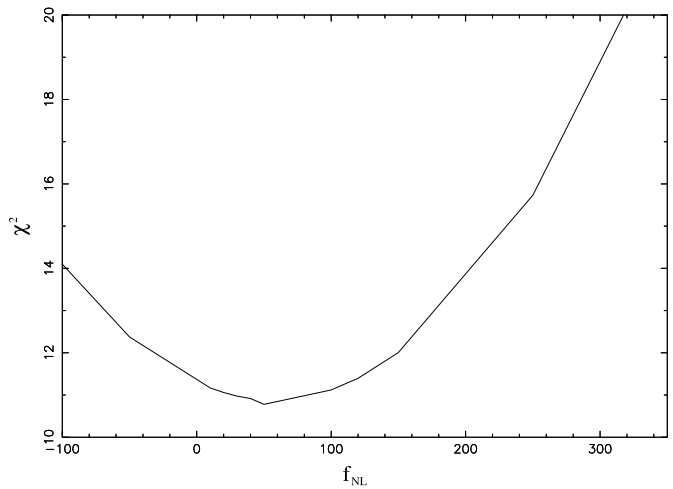

FIG. 9.—χ^2 against $f_{\rm NL}$ obtained using *WMAP* data, where $f_{\rm NL}$ is thus estimated to be 50 ± 80 at 68% confidence, and the 95% and 99% upper limits are 220 and 280, respectively.

of 50, 100, and 150. The simulations include noise and window functions of the first-year *WMAP* data in the same way as described earlier.

A plot of χ^2 values obtained using data for different $f_{\rm NL}$ is shown in Figure 9.[5] Therefore, $f_{\rm NL}$ is estimated to be 50 ± 80 at 68% confidence, with 95% and 99% upper limits of 220 and 280, respectively.

The limits on $f_{\rm NL}$ can also be checked using the Fisher discriminant function (Barreiro & Hobson 2001; Cayón et al. 2003). An optimal linear function of the measured variables (here, the skewness of the wavelet coefficients on 15 different scales) is

$$t(\boldsymbol{x}) = (\boldsymbol{\mu}_0 - \boldsymbol{\mu}_1)^T \boldsymbol{W}^{-1} \boldsymbol{x}. \quad (5)$$

Here $\boldsymbol{x}$ is the 15 element vector that contains the skewness values at the different scales considered, and $t(\boldsymbol{x})$ is the Fisher discriminant function, which optimally puts together information contained in $\boldsymbol{x}$ in the sense of maximizing the difference between the expected mean values of t from Gaussian and non-Gaussian models and minimizing their dispersions. The matrix $\boldsymbol{W} = \boldsymbol{V}_0 + \boldsymbol{V}_1$, the sum of the covariance matrices of the test statistic in the Gaussian (subscript 0) and non-Gaussian (subscript 1) cases, and μ_0 and μ_1 are vectors containing the mean values of the test statistic in the Gaussian and non-Gaussian cases, respectively.

Thus, for different $f_{\rm NL}$, the Fisher discriminant function is found for each of the Gaussian realizations, for each of the non-Gaussian realizations, and for the data. The probability that the data are drawn from one or the other hypothesis can then be estimated. Looking at the fraction of non-Gaussian simulations that have values that are larger (for positive $f_{\rm NL}$) and smaller (for negative $f_{\rm NL}$), we deduce limits on $f_{\rm NL}$ similar to those derived above using the χ^2 test. Barreiro & Hobson (2001) found that the Fisher discriminant can do better than using χ^2 values at distinguishing between Gaussian and non-Gaussian hypotheses. However, since we have only 300 non-Gaussian simulations to obtain the fractions from, the accuracy of this method is not better, but just comparable, for our case here. Figure 10 shows the histograms of the Fisher discriminants of 1000 Gaussian realizations, of 300 non-Gaussian realizations, and of the data for $f_{\rm NL}$ values of 120 and 250; these values are close to the 1 and 2 σ limits derived using the χ^2 values above. For these values of $f_{\rm NL}$, 0.74 and 0.95 of the non-Gaussian simulations, respectively, have larger values for the Fisher discriminant than the data. We see that results obtained using Fisher discriminants are consistent with those obtained using the χ^2 test.[6]

We note that if the skewness signal in the data showed deviations from Gaussianity on particular scales, then this method involving wavelet transforms could be used to obtain scale-dependent constraints on $f_{\rm NL}$. However, the skewness signal in the data is very consistent with Gaussianity. The skewness spectrum of Figure 1*a* really flattens out with the use of another mask, such as the one used in Figure 4*a*, while the kurtosis signal remains unchanged. Thus, we do not believe there is reason to distinguish between different scales in obtaining constraints on $f_{\rm NL}$ here. In addition, the constraints on $f_{\rm NL}$ derived here could possibly be made more stringent if we used the above method on the Wiener-filtered map of primordial perturbations, as discussed in Komatsu et al. (2003a). We will explore this in a future paper.

5. CONCLUSIONS

We have analyzed the first-year *WMAP* data using a spherical Mexican hat wavelet. We detect non-Gaussianity at ~99% significance, consistent with that reported by Vielva et al. (2004). This detection corresponds to a positive kurtosis

[5] This χ^2 plot is obtained from using a diagonal covariance matrix of the skewness values on different scales. One thousand Gaussian simulations may be too small to obtain convergence for the off-diagonal elements of the covariance matrix. However, as also discussed in Eriksen et al. (2004a), if we are interested in obtaining the probability of the data given the Gaussian hypothesis, as was done in § 2 of this paper, valid results can be obtained if we compute the χ^2 of both the data and the Gaussian realizations using a diagonal covariance matrix. Similarly, a diagonal covariance matrix can be used here when we are interested in the relative change in χ^2 with respect to a parameter, although we do obtain consistent limits on $f_{\rm NL}$ on using the full covariance matrix.

[6] For reference, the statistical power of the Fisher discriminant test, as defined, for example, in Barreiro & Hobson (2001), is found to be 0.28, 0.66, 0.88, and 0.98 for $f_{\rm NL}$ values of 120, 250, 350, and 500, respectively, at the 95% confidence level. These fractions are approximate, as they have been obtained using only 300 non-Gaussian realizations.

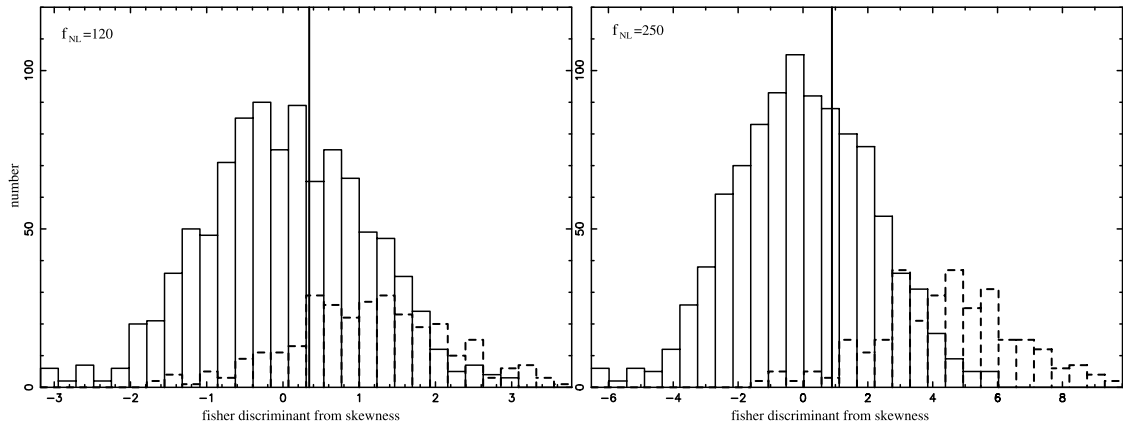

FIG. 10.—Histograms of Fisher discriminant values from 1000 Gaussian simulations (*solid lines*), 300 non-Gaussian simulations (*dashed lines*), and the data (*vertical line*), for $f_{\rm NL}$ values of 120 and 250. These values are near the 1 and 2 σ limits on $f_{\rm NL}$.

and to the presence of a larger than expected number of cold pixels (wavelet coefficients) in the southern Galactic hemisphere on scales of 3°–5°.

We have tested for changes in the significance of the signal with the type of mask used. The signal is found to be robust and is found in the ILC map as well. We have also compared confidence contours obtained for the kurtosis spectra using the full-noise simulation maps provided by the *WMAP* team, containing $1/f$ noise and other effects from data processing, to those obtained from using simulations that contain just white noise. We find very good agreement.

We have also applied another test statistic, the scale-scale correlations between wavelet coefficients. Significant scale-scale correlations are seen among the coefficients over the range of scales that indicate the above non-Gaussianity.

We then use the skewness statistic on the different scales to place constraints on the nonlinear coupling parameter $f_{\rm NL}$, the motivation being to see how much non-Gaussianity of this particular form is allowed by the current data. It is also a way to compare the sensitivity of different test statistics to this parameter. The constraints obtained are closely consistent with those obtained by Komatsu et al. (2003b) using the cubic statistic and Minkowski functionals on the same data. The constraints on $f_{\rm NL}$ derived here could possibly be made more stringent if we used spherical wavelets on the Wiener-filtered map of primordial perturbations, as discussed in Komatsu et al. (2003a). We will explore this in a future paper. The kurtosis statistic is not sensitive to this form of non-Gaussianity. We will present constraints on other forms of non-Gaussianity implied by the kurtosis statistic of the *WMAP* data elsewhere.

We acknowledge the use of non-Gaussian simulations from Komatsu et al. (2003b; a *WMAP* paper). It is a pleasure to thank Eiichiro Komatsu for beneficial discussions and comments on the manuscript. P. M. thanks Gang Chen for useful discussions. We acknowledge the use of the OSCER supercomputing facilities at the University of Oklahoma. This work is supported in part by NSF CAREER grant AST 00-94335.

APPENDIX

The SMHW transform has been used previously by Cayón et al. (2001, 2003), Martínez-González et al. (2002), and V04. It is a continuous and symmetric wavelet that in the small-angle limit corresponds to the Euclidean Mexican hat wavelet. The wavelet is given by

$$\Psi_R(y) = \frac{1}{\sqrt{2\pi}N(R)}\left[1+\left(\frac{y}{2}\right)^2\right]^2\left[2-\left(\frac{y}{R}\right)^2\right]e^{-y^2/2R^2}, \tag{6}$$

where R is the scale, $N(R)$ is a normalization constant given by

$$N(R) = R\left(1+\frac{R^2}{2}+\frac{R^4}{4}\right)^{1/2}, \tag{7}$$

and $y = 2\tan\theta/2$, for polar angle θ. Therefore, ψ is isotropic so that when centered on the north pole it is independent of the polar coordinate ϕ. Thus, moving any point to the north pole, the wavelet coefficients

$$w(R,\ x' : (\theta',\ \phi')) = \int dx\, T(x)\Psi_R(|x'-x|) \tag{8}$$

can be obtained via a convolution of the sky map $T(x : (\theta,\ \phi))$ with the wavelet function. This is easily performed in spherical harmonic space.

REFERENCES

Aghanim, N., Kunz, M., Castro, P. G., & Forni, O. 2003, A&A, 406, 797
Barreiro, R. B., & Hobson, M. P. 2001, MNRAS, 327, 813
Barreiro, R. B., Hobson, M. P., Lasenby, A. N., Banday, A. J., Górski, K. M., & Hinshaw, G. 2000, MNRAS, 318, 475
Bennett, C. L., et al. 2003, ApJS, 148, 97
Cabella, P., Hansen, F., Marinucci, D., Pagano, D., & Vittorio, N. 2004, Phys. Rev. D, 69, 063007
Cayón, L., Martínez-González, E., Argüeso, F., Banday, A. J., & Górski, K. M. 2003, MNRAS, 339, 1189
Cayón, L., Sanz, J. L., Martínez-González, E., Banday, A. J., Argüeso, F., Gallegos, J. E., Górski, K. M., & Hinshaw, G. 2001, MNRAS, 326, 1243
Chiang, L.-Y., Naselsky, P. D., Verkhodanov, O. V., & Way, M. J. 2003, ApJ, 590, L65
Colley, W. N., & Gott, J. R. 2003, MNRAS, 344, 686
Copi, C. J., Huterer, D., & Starkman, G. D. 2004, Phys. Rev. D, in press (astro-ph/0310511)
Eriksen, H. K., Hansen, F. K., Banday, A. J., Górski, K. M., & Lilje, P. B. 2004a, ApJ, 605, 14
Eriksen, H. K., Novikov, D. I., Lilje, P. B., Banday, A. J., & Górski, K. M. 2004b, ApJ, 612, 64
Gaztanaga, E., & Wagg, J. 2003, Phys. Rev. D, 68, 021302
Gurzadyan, V. G., et al. 2004, preprint (astro-ph/0402399)
Hansen, F. K., Cabella, P., Marinucci, D., & Vittorio, N. 2004, ApJ, 607, L67
Hinshaw, G., et al. 2003, ApJS, 148, 135
Hobson, M. P., Jones, A. W., & Lasenby, A. N. 1999, MNRAS, 309, 125
Komatsu, E., Spergel, D. N., & Wandelt, B. D. 2003a, preprint (astro-ph/0305189)
Komatsu, E., Wandelt, B. D., Spergel, D. N., Banday, A. J., & Górski, K. M. 2002, ApJ, 566, 19
Komatsu, E., et al. 2003b, ApJS, 148, 119
Martínez-González, E., Gallegos, J. E., Argüeso, F., Cayón, L., & Sanz, J. L. 2002, MNRAS, 336, 22
Mukherjee, P., Hobson, M. P., & Lasenby, A. N. 2000, MNRAS, 318, 1157
Pando, J., Valls-Gabaud, D., & Fang, L.-Z. 1999, Phys. Rev. Lett., 81, 4568
Park, C.-G. 2004, MNRAS, 349, 313
Santos, M. G., et al. 2003, MNRAS, 341, 623
Smith, S., et al. 2004, MNRAS, in press (astro-ph/0401618)
Spergel, D. N., et al. 2003, ApJS, 148, 175
Starck, J. L., Aghanim, N., & Forni, O. 2004, A&A, 416, 9
Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., & Cayón, L. 2004, ApJ, 609, 22 (V04)